From Contextual Search to Data Storytelling: Building Semantic Competence on Multiple AI Substrates

Larry Birnbaum

Northwestern University Department of Computer Science Evanston, Illinois USA

Jio Institute AI & Data Science Program Navi Mumbai, India

Semantic Competence

- Or rather: **semantic** <u>and</u> **pragmatic** competence.
- By this I mean simply, the ability to extract or convey something meaningful and useful concerning the content of natural language in context of use.
- Common sense inference and question-answering — whether explicit or implicit — are key components of these capabilities.

Application Areas and AI Substrates

- Intelligent information systems ⇔ IR
- Computational journalism ⇐ Text analytics
- Data storytelling \(\Leftarrow Specialized inference models / limited context systems \)
- Conversational interaction <>> Neural language models

Intelligent Information Systems

• Building intelligent systems on top of Information Retrieval (IR) as the substrate.

The Eye of the Needle

Advanced Search Language Tools

First Solution: Watson



Watson: Mechanism

- Context analysis
- Query formation
- Query management
- Result filtering and ranking
- Presentation

• Fast, automatic, better than human performance

Determining Relevance

- We implicitly utilized the standard IR paradigm in which similarity *stands as a proxy* for relevance.
- But the results, while "on point," were often seen by users as just "more of the same."

Similarity != Relevance

- The most similar possible document to a document that you have in your hands is... another copy of that document.
- What makes a document relevant is that it is similar in certain respects and *dissimilar* in certain other respects.
- Watson relied on noise and the size of the internet to achieve the variation necessary to provide truly relevant documents.
- Can we do better?

Relevance = Editorial Judgment

- Determining what information people see and in what form and order — is an *editorial judgment*.
- This editorial judgment should be *explicit* and *visible* to both designers/engineers and users.
- This means a *deliberate* and *understandable* mechanism.
 - Which can only? be realized by a *semantic* mechanism.
- Specific *dimensions of similarity/dissimilarity* provide useful information to the user based on his/her context of activity and use—i.e., answer questions the user might plausibly have in that

context.

Beyond Broadcast



\equiv delish	DINNERS	RECIPES	FOOD NEWS	COCKTAILS	TIPS & TOOLS
	YIELDS: 8 serving(s)	PREP	TIME: TO mins 1	DTAL TIME:	CAL/SERV: 533

Ingredients

3/4 lb. lasagna noodles 1 tsp. extra-virgin olive oil, plus more for drizzling 2 lb. ground beef 4 cloves garlic, minced 2 tsp. dried oregano Kosher salt Freshly ground black pepper 2 (32-Oz.) jars marinara 16 oz. whole milk ricotta 1/2 c. freshly grated Parmesan, divided 1/4 c. chopped parsley, plus more for garnish 1 large egg 2 lb. sliced mozzarella

Directions

Step 1

Preheat oven to 375°. In a large pot of salted boiling water, cook pasta according to package directions until al dente, less 2 minutes. Drain and drizzle a bit of olive oil to prevent noodles from sticking together.

SAVE TO MY RECIPES

Step 2

Meanwhile, in a large pot over medium-high heat, heat oil. Cook ground beef until no longer pink, breaking up with a wooden spoon. Remove from heat and drain fat. Return beef to skillet and add garlic and oregano and cook, stirring, for 1 minute. Season with salt and pepper, then add marinara and stir until warmed through.

Step 3

Combine ricotta, 1/4 cup Parmesan, parsley, and egg in a large mixing bowl and season with salt and pepper. Set aside.

Relevant Semantic Dimensions

...in the domain of food recipes:

- Low fat
- Low salt
- Low cal
- Vegetarian / vegan
- Quick and easy
- Etc.

Compare & Contrast



Two Comparable Stories

. . .

Oracle tried to buy open-source **MySQL**

SAN FRANCISCO--Oracle tried to acquire open-source database maker MySQL, an indication of the profound changes the software giant is willing to make as it adapts to the increasingly significant collaborative programming philosophy.

MySQL Chief Executive Marten Mickos confirmed the **acquisition** attempt in an interview at the Open Source Business Conference here but wouldn't provide **details** such as when the approach was made or how much money **Oracle** offered.

. . .

. . .

Oracle didn't immediately comment on the **acquisition** offer.

Though it is increasingly diversified, Oracle's primary business is selling its own proprietary database software. MySQL, in contrast, is a leader among several companies trying to commercialize rival open-source **products**.

IBM Expands Paid Open Source Strategy

UPDATED: **IBM** is making a bid on professional **open source** with the **acquisition** of privately held **Gluecode**, officials announced **Tuesday**.

Officials did not discuss financial and operational **details** of the merger, the first **acquisition** made by **Big Blue** of an **open source** company.

Gluecode's operations will be assimilated into IBM's **software** group and expand the company's WebSphere application integration middleware product line.

Officials plan to offer customers and business partners Gluecode's application server software and sell software and support services on top of the offering, as well as let customers upgrade to IBM WebSphere products.

from internetnews.com

from CNET News.com

Another Example



LocalSavvy



It described the allegation as "psychological warfare" calculated by Tebran's enemies - an Done Assassins plot to kill Putin in Iran: Interfax 15 Oct 2007 - 25 weeks ago (A)

LocalSavvy Cont'd



Blog Search: Spectrum

😻 Expert Bloggers - Mozilla Firefox		
Eile Edit View History Bookmarks Tools Help		\diamond
🖼 🤄 - 🖻 - 🎯 👔 🗋 http://prism.cs.northwestern.edu/expertBlogger/	🔹 🕨 Wikipedia (English)	Q
Bispectrum for the second seco		
arts _business _education _health _law _politics _religion _technology Search for blogs "opistomic" dimon	sion in which t	
query keywords user is interested		

Spectrum Results

Search results for EU Intel

T (1 00 1 1 0 1

business 13 law 4 politics 10 technology 23

1.01.46 0.30

- <u>Record EU Fine for Intel, Deconstructing Intel Response | BNET ...</u>
 From BNET Technology http://industry.bnet.com/technology/ 13 May 2009
 BNET Technology provides daily industry news coverage and insights for managers and executives about all aspects of the high-tech industry.
- <u>AMD</u>: EU ruling against Intel a step towards establishing a truly ...
 From PCLaunches.com http://www.pclaunches.com/ 47 minutes ago
 Here's official word from AMD regarding EU's ruling against Intel: EU ruling against Intel is "a step towards establishing a truly competitive market." "With this ruling, the industry will benefit from an end to Intel's ...
- Intel officially fined a record \$1.45 Billion by European Union ...
 From TheNextWeb.com http://thenextweb.com/ 9 hours ago
 The commission told Intel to immediately stop some sales practices in Europe, though it wouldn't say what those were. Intel said it was ?mystified? about what it was supposed to change but would comply while it appeals the fine. ...

From technologist and IT entrepreneurs Focus on the impact on the IT industry

Spectrum Results Cont'd

Search results for EU Intel

business 13 law 4 politics 10 technology 23

- TRUTH ON THE MARKET » The EU Intel Decision, Error Costs, and What ... From TRUTH ON THE MARKET - http://www.truthonthemarket.com/ 1 hour ago
 Of course, the EU approach does not make room for such safe harbors. Not even close. To describe the EU approach to Intel's loyalty rebates as either remotely ?effects-based? or ?evidence-based? would strip both those terms of any ...
- <u>Antitrust Week Continues: EU Slams Intel With \$1.45b Fine Law ...</u> From WSJ.com: Law Blog - http://blogs.wsj.com/law/ 13 May 2009
 European Union regulators hit Intel Corp. with a \$1.45 billion fine -- the largest ever assessed for monopoly abuse -- and called for changes in the way the
 US company sells the microprocessors at the heart of most of the world's ...
- <u>Tech Law Prof Blog: EU Fines Intel Over Uncompetitive Practices in ...</u> From Tech Law Prof Blog - http://lawprofessors.typepad.com/tech_law_prof/ 13 May 2009
 Intel is fined a whopping \$1.45 Billion for market practices the European Union considers anti-competitive. Specifically, Intel used rebates for its processors with terms that made its customers either reject or restrict the use of AMD ...

From lawyers and law school professors Focus on the legal issues of the case

Tell Me More



Tell Me More: Actors

BP's test of newly installed cap is put off

Chicago Tribune | 2010-07-14 08:36:28 |

Reporting from Atlanta and New Orleans-- The head of the federal response to the gulf oil spill announced Tuesday night that the government and BP would delay a crucial test to determine whether it was safe to use a newly applied sealing cap to finally shut down the runaway well.

It is hoped that the test will determine whether there are any serious leaks plaguing the 13,000-foot hole in the earth apart, of course, from the notorious leak at the top.

The test was originally scheduled to take place Tuesday evening. Now it appears that it will not take place until late Wednesday at the earliest.

In a statement on the government's spill response website, former Coast Guard Adm. Thad Allen said the timetable changed after a discussion he had with Energy Secretary Steven Chu, BP officials, geologists and other experts.

"As a result of these discussions, we decided that the process may benefit from additional analysis that will be performed tonight and tomorrow," Allen said in a statement.

A representative from the spill response's information center could not elaborate on the reasons for the delay, and BP and Department of Homeland Security officials did not respond to calls late Tuesday.

ADDITIONAL ACTORS

Stuff.co.nz

<u>BP puts oil cap test on hold - americas - world |</u> <u>Stuff.co.nz</u>

BP divested a small non-core downstream asset on Tuesday, selling Magellan Midstream Partners L.P., its crude oil storage tanks in Cushing, Oklahoma and related petroleum pipelines for \$289 million.

Channel 4 News

<u>BP successfully places cap on oil well - Channel 4</u> <u>News</u>

Chief Executive of American Petroleum Institute, Jack Gerard said: "It is unnecessary and shortsighted to shut down a major part of the nation's energy lifeline while working to enhance offshore safety.

MORE



ADDITIONAL QUOTES

Tell Me More: Data

BP's test of newly installed cap is put off

Chicago Tribune | 2010-07-14 08:36:28 |

Reporting from Atlanta and New Orleans-- The head of the federal response to the gulf oil spill announced Tuesday night that the government and BP would delay a crucial test to determine whether it was safe to use a newly applied sealing cap to finally shut down the runaway well.

It is hoped that the test will determine whether there are any serious leaks plaguing the 13,000-foot hole in the earth apart, of course, from the notorious leak at the top.

The test was originally scheduled to take place Tuesday evening. Now it appears that it will not take place until late Wednesday at the earliest.

In a statement on the government's spill response website, former Coast Guard Adm. Thad Allen said the timetable changed after a discussion he had with Energy Secretary Steven Chu, BP officials, geologists and other experts.

"As a result of these discussions, we decided that the process may benefit from additional analysis that will be performed tonight and tomorrow," Allen said in a statement.

A representative from the spill response's information center could not elaborate on the reasons for the delay, and BP and Department of Homeland Security officials did not respond to calls late Tuesday.

ADDITIONAL ACTORS

ADDITIONAL NUMBERS

Commodity Surge (blog) <u>Commodity Surge:: Why New BP (NYSE:BP)</u> Containment Cap Takes Time to Test

What scientists want to see is the pressure readings to come in at around 8,000 to **9,000** pounds per square inch. That would indicate there are no other leaks coming from the oil well. If it's under about 6,000, they say it would mean there are leaks that hadn't been identified yet.

The Associated Press

The Associated Press: BP: New cap will be on busted well later today

Ultimately, the plan is to have four vessels collecting oil from the leak with a combined capacity of about 2.5 million to **3.4** million gallons a day - enough to capture all the oil leaking, if federal estimates are right. Getting all the vessels on the task will take about two to three weeks.

MORE

Tell Me More: Quotes

BP's test of newly installed cap is put off

Chicago Tribune | 2010-07-14 08:36:28 |

Reporting from Atlanta and New Orleans-- The head of the federal response to the gulf oil spill announced Tuesday night that the government and BP would delay a crucial test to determine whether it was safe to use a newly applied sealing cap to finally shut down the runaway well.

It is hoped that the test will determine whether there are any serious leaks plaguing the 13,000-foot hole in the earth apart, of course, from the notorious leak at the top.

The test was originally scheduled to take place Tuesday evening. Now it appears that it will not take place until late Wednesday at the earliest.

In a statement on the government's spill response website, former Coast Guard Adm. Thad Allen said the timetable changed after a discussion he had with Energy Secretary Steven Chu, BP officials, geologists and other experts.

"As a result of these discussions, we decided that the process may benefit from additional analysis that will be performed tonight and tomorrow," Allen said in a statement.

A representative from the spill response's information center could not elaborate on the reasons for the delay, and BP and Department of Homeland Security officials did not respond to calls late Tuesday.

ADDITIONAL ACTORS

ADDITIONAL NUMBERS

ADDITIONAL QUOTES

The Associated Press <u>The Associated Press: BP: New cap will be on</u> busted well later today

"The hope is that we can slowly turn off the valves, close the capping completely and then test pressure to see how the well is performing," Thad Allen, the government's point man on the disaster, said on CBS's "The Early Show."

Stuff.co.nz

<u>BP puts oil cap test on hold - americas - world |</u> <u>Stuff.co.nz</u>

"We are in discussions with a number of companies about a number of assets," BP spokeswoman Sheila Williams said in London, declining to give details. "Talks are going well."

MORE

The Context Project

A platform and toolset to support:

- Content Analysis
- Source Determination
- Query Formation
- Query Management
- Result Ranking and Extraction
- Presentation

TweetTalk

RT @aasif: Glad to see the protests in France against terrorist attacks. But Boko Haram killed almost 2000 pple on Friday,... http://t.co/P...

- Orwell (@_Orwell) Mon Jan 12 05:03:17 +0000 2015

France: 3.7m people rally for unity against terror after Paris attack http://t.co/aa8DJs6sM5

- Rahul Ranjan (@rahulranjan535) Mon Jan 12 05:03:16 +0000 2015

#newsDOTpk RT dunyanetwork: Standing ovation at Golden Globes for Paris attacks message ... http://t.co/Q5QOh7Ob3E http://t.co/cfKcUFNXsl

- News.pk (@newsDotPK) Mon Jan 12 05:03:13 +0000 2015

RT @DaisyKhan: Worth reading! @TIME: Kareem Abdul-Jabbar: "These terrorist attacks are not about religion" http://t.co/VtmvVkd8vD"@kaj33

- salman ahmad (@sufisal) Mon Jan 12 05:03:13 +0000 2015 TweetTalk | Northwestern University InfoLab

readdit

	-	Keywords		
16 January 2015 Last updated at 12:52 ET		police, Paris, attacks, raids, Friday, France, people, gunmen, Belgium, Mr Kerry		
		Related Reddit Threads		
Paris attacks: Twelve suspects he	elc	Twelve suspects are being held by police in the Paris region over last week's Muslim terrorist attacks in the French capital that killed 17 people		
overnight		Twelve held over Paris attacks		
		'Ten questioned' over Paris attacks		
US Secretary of State John Kerry visited the Charlie Hebdo offices, as the BBC's Chris Eakin reports		Police Chief In Charge of Paris Attacks Commits Suicide		
		Australian government exploits Paris attacks -"Abbott immediately seized upon		
Twelve suspects are being held by police in the Paris region over last week's attacks in the French capital that killed 17 people.	Pa	the attack in Paris to foreshadow further boosting the vast powers of Australia's police-intelligence apparatus and expanding the country's frontline participation in the US-led war in Iraq and Syria."		
They are being questioned about "possible logistical support", such as weapons or vehicles, they could have given the three gunmen, police say.		Paris Shooting Megathread - II		
		Operation Gladio, Alive and Well? Paris Attackers Funded by Pentagon Dinner Guest, and 5 Other "Coincidences"		
Police conducted raids in five towns in the Paris region.	Thr	Suspects in Paris attack were on police radar for years		
Fr		Brother of Muslim police officer slain during Paris terrorist attack says terrorists		
Last week's violence began with an attack by two gunmen on Chanie Hebdo magazine, which left 12 people dead		aren't Muslim. Let's see how /r/worldnews handles that.		
		Paris Shooting Megathread		
A day later - on 8 January - a policewoman was shot dead by gunman		Eric Bolling: we need to "overmilitarize" the police after the attack in Paris		
Amedy Coulibaly while responding to a traffic accident in the south of Paris.		driven to Kosher supermarket by an accomplice		
		Attacks in France obvious false flag. All that planning and they leave their ID's in		
Four more people who were held hostage by Coulibaly were killed on 9		their decoy car.		
January when police stormed a kosher supermarket. Coulibaly was shot dead by police. Police also killed the two gunmen responsible for the		Readdit Northwestern University InfoLab		
Charlie Hebdo attack - Said and Cherif Kouachi - outside a warehouse				
near Paris earlier that day.				
Three members of the magazine's staff were buried on Friday, including				
the editor Stephane Charbonnier, also known as Charb. Bagpipers played				
Amazing Grace at his funeral in the Pontoise district of Paris.				

readdit cont'd

MY SUBREDDITS 🛫 🛛 FRONT - ALL - RANDOM PHOTOSHOPBATTLES - FITNESS - TELEVISION - NOSLEEP - MUSIC - SHOWERTHOUGHTS - DATAISBEAUTIFUL - NEWS -	- DOCUMENTARIES - ASKREDDIT - ASKSCIENCE - ART - MORE »	
CHARLIE EUROPE comments related other discussions (1)	want to join? sign in or create an account in seconds English	
Paris Shooting Megathread (self.europe) submitted 6 days ago * by BuropeMods		
Key Facts	///Europe	
 12 dead and 5 injured after gunmen storm offices of French satirical magazine. 		
3 attackers on the run after gunright with police.		
Highest terror alert in Paris at the moment.		
 Map of incident, thanks to /u/KetchupTubeAble19 		
Live Updates	Q search Q	
BBC News	this post was submitted on 07 Jan 2015	
France24	2,179 points (93% upvoted)	
Le Monde		
AFP Livestream from crime scene	shordlink: http://redd.it/2rmi3x	
To prevent the front page being full of the same story, we are creating this thread to collate all the stories and confirmed infomation.	username password	
tag /u/EuropeMods to include stories in the title	remember me reset password login	
English		
 BBC News - 'Shots fired' at HQ of French satirical magazine Charlie Hebdo (Previous reddit discussion - >500 comments) 	Submit a new link	
Daily Telegraph - 'Shooting reported at Paris magazine Charlie Hebdo'	Submit a new text post	
Twitter account of Charlie Hebdo showing final tweet sent	025.029	
 SFW picture of attackers moments before they execute a wounded policeman. Some may find this disturbing 	subscribe 215.380 Europeans	
 Image of forensic teams examining a car used in by attackers. 	O 255 Europeans	
CBS News - Paris newspaper attacked - In Pictures		
Statement from the US President.	FILTER RUSSIA, UKRAINE, NATO	
Image of bulletholes in police car.	NEW TO REDDIT? CLICK HERE!	
The Telegraph - 'Terrorists shouted they were from al-Qaeda in the Yemen before Charlie Hebdo		

Stakeholder Tweetback

STAKEHOLDERTWEETBACK

Tweets

@BarackObama

The senators who voted against background checks for gun sales want you to forget. Show them you won't: http://t.co/Yjzjyh0dL7

This is big: @OFA volunteers are about to deliver 1.4 million signatures to Congress demanding expanded background checks for gun sales.

On Thursday, @OFA will deliver to Congress the signatures of 1.4 million Americans who want expanded background checks for gun sales.

More than 1.4 million Americans signed our petition demanding expanded background checks for gun sales—by far the most in @OFA history.

Tell Congress now is the time for background checks for gun sales: http://t.co/w2oHKvXokq Add your name by midnight tonight.

@SenMarkPryor

@sherrilee7 Yes. This bill doesn't apply to hunting licenses. This is about the universal background check requirements in the current bill.

I've gotten a lot of questions about NYC Mayor gun ad. My response? I don't take gun advice from the Mayor of NYC. I listen to Arkansans.

Imposters r calling homes saying they're w my office & asking ab # of guns fams own. Dirty political trick. Im strong 2 amdt guy & gun owner

Nat'l Assoc 4 Gun Rights petition is blatantly false. As hunter & gun owner, I'm a strong supporter of 2 amdt rights & opponent of AW ban.

Mayors Against Illegal Guns targets first Democrat - latimes.com

In an ad that launched Friday, Sen. Mark Pryor (D-Ark.) is implored by a witness to the murder of state Democratic Party Chair Bill Gwatney in 2008 to revisit his vote on the failed bipartisan effort to expand background checks on gun purchases.

"The pain of that experience will always be a part of me," Angela Bradford-Barnes says in the ad. "Like 84% of Arkansans I support universal background checks. Thus, I was heart broken to see that Senator Pryor opposed the bipartisan bill because it will save lives.

All 23 of President Obama's gun policy proposals

The Mayors Against Illegal Guns ad is its first targeting a Democrat, following earlier ads launched against Sens. Kelly Ayotte (R-N.H.) and Jeff Flake (R-Ariz.).

Pryor, in a statement Friday, condemned the ad as "disgusting," accusing the group of politicizing Gwatney's death.

"New York City Mayor Mike Bloomberg didn't know Bill Gwatney. I knew Bill Gwatney. He was my friend and he was killed by someone with severe mental health issues," Pryor said. The mayor's bill would have done nothing to prevent his death because it falls to adequately address the real issue and common thread in all of these shootings — mental health."

Pryor hasn't shut the door completely on new gun control measures, saying earlier this month that he would consider a combination of the defeated proposal from Sens. Joe Manchin (D-W.V.) and Pat Toomey (R-Pa.) and an alternative from Sens. Chuck Grassiey (R-Iowa) and Ted Cruz (R-Texas).

PHOTOS: The debate over gun control

Under the Grassley-Cruz plan, straw purchasing and trafficking of guns would be criminalized and school security as well as efforts to keep guns out of the hands of the mentally ill would be expanded. Background checks, however, would remain largely untouched.

Pryor, up for reelection in 2014, is already facing attacks from the right, with Club for Growth Action and the Senate Conservatives Fund already launching ads against his candidacy. And in addition to Mayors Against Illegal Guns, the pro-Obama group Organizing for Action has also announced Pryor as a target for his vote on background checks.

Follow Politics Now on Twitter and Facebook

Search

Article URL:

Go

Computational Journalism

- Building computational journalism systems on top of text analytics using standard machine learning methods.
- Playing to the machine's strengths treat natural language more as data (and less as a medium of communication).

Tweetcast 2012



Tweetcast 2016



Tweetcast Norway

Knight Lab | knightlab.edu | @knightlab | @tweetcast



TweetCast your vote

😏 Sign in with Twitter

TweetCast predicts which political party you are most likely to vote for in the 2013 Norwegian parliamentary election based on an analysis of your Twitter activity.

Our technology examines the information in your tweets, who you are friends with, and your profile description to form a prediction. Below we have listed the most predictive attributes for each of the major political parties participating in this election.

BookRecs



BookRecs

- BookRecs recommends books to you based on your social media data.
- > 60K people have tried the system.
- About 15% have clicked through to look at a book more closely.
LocalRecs



LocalRecs

- LocalRecs recommends restaurants, bars, cultural venues, and shops to you based on your tweets.
- For popular establishments, the same as Yelp.
- For less popular places i.e., niche interests — *better*.

Student Club Recommender

- Hillary Clinton
 - NU Democrats, College Republicans, Northwestern
 Political Union
- Julia-Louis Dreyfus
 - NU Democrats, Jewish Theatre Ensemble, Northwestern Flipside [satirical magazine], Lipstick Theatre [women's theater group], Dolphin Show [musical theater group], Student Alumni Alliance
- Bill Gates
 - Unite for Sight [global health], College Republicans, Northwestern Crew, Engineers for a Sustainable World, Alpha Chi Omega [sorority], Design for America

Building a Platform

- When you've built a few applications that look pretty similar, the next step is to build a *platform* that will make it much easier to create applications of that sort.
- You input the district and the candidates, and the system builds a Tweetcast application and web site automatically.

Predictive Politics

predictive politics.



where what you say, says you how you'll vote.

predict a vote

build new election model

Narrative Analysis and Computational Journalism

- Our goal in this work is to analyze news stories in functional, pragmatic, or thematic terms that focus on narrative.
- We aim for factors that are "orthogonal" to highly studied dimensions of text such as topic detection/modeling or sentiment analysis.
- The ultimate aim is to give readers better understanding and control of the information they receive.

Categorizing Election News

- Election news stories fall into a number of different genres or frames, in general terms:
 - Policy
 - Candidate (biography / character)
 - "Horse race"
- Horse race stories, focused on polling, money, strategy, "ground game," etc.—i.e., whose "ahead"—constitute a very large percentage of election stories in the US.
- There is some evidence that these stories promote a sense of cynicism among voters.

Detecting Horse Race Stories

- We tried a number of approaches using a simple bag of words representation of election news stories.
- The best results were obtained using hand-built term sets!
 - Precision = .7
 - Recall = .69
- More sophisticated representations / algorithms could probably improve on this.

Publication "Report Cards"

2017 Alabama Senate Special Election

	% Horse	· · ·
Publication	Race	n
patch.com	14.3	7
CNN	50.0	6
Washington Times	50.0	16
Seattle Times	54.5	11
Daily Mail	55.6	9
Blog	60.0	5
The Oklahoman	60.0	5
Chicago Tribune	62.5	8
Fox News	62.5	8
Washington Examiner	63.6	11
newsmax.com	66.7	6
New York Times	66.7	6
Bleeding Heartland	80.0	5
Boston Herald	80.0	5
CBS News	80.0	5
LA Times	80.0	5
Times-		
Picayune/nola.com	80.0	5
syracuse.com	80.0	5
Tennessean	83.3	6
yahoo.com	83.3	6
Politico	84.6	13
MSN	88.9	9
Breitbart	90.0	10
Washington Post	93.8	32

Offering Readers an Alternative



Heroes, Villains, and Victims



Heroes, Villains, Victims

- Again using simple classifiers and a bag of words representation.
- Task is not easy but performance is mediocre and needs considerable improvement:
 - Villain F1 = .4
 - Victim F1 = .43
 - Hero F1 = .55

Telling Stories at Scale

• There's no point in gathering and analyzing all the data we are gathering and analyzing if we can't convey the results to people in a way that they can understand and take action on!

Data Storytelling

- Automatic generation of natural language stories based on data, and aimed at conveying key insights from those data.
- More semantics more focus on language as a medium of communication.
- To make it work, you need to find a powerful and useful "sweet spot" in terms of task / domain / context in which you can engineer a solution.

The Punch Line

R	o Grande Valley Kille	r Bees																								
	# PLAYERS					Ρ	GP	G	Α	PTS	+/-	PIM	PPG	PPA	SHG	i Sl	HA	GWG	SHOT	S SI	РСТ					
	71 Gendur, Dan					F	41	14	29	43	-8	32	3	14	0		0	2	186	.0)75					
	21 Klempa, Tomas					F	42	19	18	37	-3	41	4	6	1		0	4	134	1	142					
	9 Lee, Aaron					F	43	Lar	edo B	Bucks																
	10 Campos, Brandon					F	38	-	-				-	CD	-				-	-	-	CUIC	-	CHUC	CHOTC	CRET
	15 Dixon Jacon					F	40	#	PLAYE	EKS	_		P	GP	G	A	22	+/-	PIM P	PG	PPA	SHG	SHA	GWG	126	OPS
	X Ha GAME DETAIL	s																			11	0	1	1	72	.095
	11 Mikk SCORING						_			SHOTS											6	2	3	1	88	.091
	X Ma				1	2	2 3	то	TAL						1	2	3	т	TAL		3	1	1	0	78	.115
	51 Bala Rio Grande	Valley			1	C) 1		2	Rio Gran	nde Va	lley			11	4	9		24		8	0	1	0	75	.093
	5 Riou Laredo				1	1 2	2 0		3	Laredo					8	10	7		25	1_	4	2	1	2	64	.156
RIO G	irande Valley							Lare	edo										_		4	0	0	1	83	.084
# N/	AME	G	А	PTS	+/-	s	PIM	# 1	NAME	_		G	А	PTS	+/-	S	; I	PIM			3	0	0	2	117	.081
5 R	oux, Nicolas	0	1	1	-1	0	2	4 1	Marian	i, Kyle		0	0	0	+1	3	3	0			3	0	0	1	87	.103
9 Le	ee, Aaron	0	0	0	-1	1	0	5 1	Moore,	Greg		0	0	0	0	2	2	0								
	ampos, Brandon	0	0	0	-1	0	2	8	Master	s, Rusty	8	1	1	2	+2	2	2	2			_		_	_	_	_
11 M	ikkelsen, AJ	0	0	0	0	2	0	9	Sparlir	ng, Ryan	3	0	0	0	+1	0)	2								
14 Pe	arson, Zac	0	0	0	-2	5	0	10	Bolsve	rt, Erik		0	1	1	+2	2	2	0								
5 D	ixon, Jason	0	0	0	-1	1	0	12	Bovd.	Matt		0	0	0	+1	1		2								
9 Is	chi, Phillip	0	0	0	-1	2	0	13	Harcha	arik. Jose	ph	0	1	1	+1	4	1	0								
21 K	empa, Tomas	1	1	2	-1	1	0	15	Pierce	Matt R		1	0	1	+1	3	3	0								
24 F	ontaine. Charles	0	0	0	-1	0	2	17 1	Inman	David		1	1	2	+3	3	3	2								
26 W	ilson. Mike	0	0	0	-1	0	0	19	Styffe	Justin		0	0	0	0	0)	0						_		
8 B	arr. Tyler	0	0	0	0	1	2	20	Hau. A	.].		0	0	0	0	2	,	0								
10 M	urray, John	0	0	0	0	1	0	23	Trimm	Nell		0	1	1	+1	0)	2								
4 R	nders. Wylie	0	0	0	0	0	0	27	Stitt	Tyler		0	0	0	+1	0)	0								
1 B	alan, Scott	0	0	0	0	2	0	28	Rivet	Adam		0	1	1	+1	3	1	2								
5 11	neker, Crain	0	1	1	-7	0	0	31	lung '	Torrie		0	0	0	0	0)	0								
57 51	artsev Kirill	0	0	0	-1	1	0	33	Ronde	au Ryan		0	0	0	0	0)	0								
1 6	endur Dan	1	1	2	-1	7	4	44	Hanson	n Matt		0	0	0	0	0)	0								
1 0	ongaard Aaron	0	0	0	-1	0			ransor	n, marc		U	U	U	U	U										
GOA		U	U	U	-1	0	U	9										3								
AME			DEC	MIN	SHO	OTS	SAVES	NAM	E				DEC	M	N S	HOTS	5 5A	VES								
0 Jol	n Murray	L	OSS	58:51	25	5	22	31 T	orrie J	ung			WIN	60:	00	24	2	22								
30 Jol 34 Wy	nn Murray rlie Rogers (b)	l	.055	58:51	25	5	22	31 T 33 R	orrie J yan Re	ung ondeau (t	b)		WIN	60:	00	24	2	22								



WO

per es rith

February 9, 2012, Hockey Recap: Rio Grande Valley rolls over Laredo, 6-3

pointstre

The Rio Grande Valley Killer Bees were firing on all cylinders against the Laredo Bucks, and when the final buzzer sounded Killer Bees emerged with a 6-3 win.

Zac Pearson was all over the ice for Rio Grande Valley, as he tallied two goals and one assist in the win. Pearson scored the first of his two goals at 5:23 into the first period to make the score 1-0 Rio Grande Valley. Brandon Campos picked up the assist. Pearson's next tally made the score 2-0 Rio Grande Valley with 12:44 left in the first period. David Marshall assisted on the tally.

^{Spe tea} ^{Cot} The Rio Grande Valley Killer Bees were ^{friduan} ^{The gar} Bucks, and when the final buzzer ^{on} sounded Killer Bees emerged with a 6-3

Rio Grande Valley additionally got points from Aaron Lee, who had one goal and one assist, Marshall, who registered one goal and two assists, and Dan Gendur, who racked up one goal and one assist. Dan Nicholls also scored for Rio Grande Valley. Others to record assists for Rio Grande Valley were AJ Mikkelsen, who had two and Adam Bartholomay and Marc-Andre Carre, who each chipped in one.

Laredo was often in penalty trouble, as it ended with six minors and one major for 17 minutes in penalty time. The leading offender was Justin Styffe, who totaled nine minutes in penalty time with two minors

Business Stories

Red Hat, Inc. Common Stock (RHT) - NYSE

🕈 Add to Portfolio 🛛 🖪 Like

< 50

42.43 + 0.68(1.58%) Jan 13, 4:02PM EST | After Hours: 42.41 + 0.02 (0.06%) Jan 13, 4:42PM EST

Prev Close:	43.11	Day's Ra	nge: 4	42.14 - 42.88	Red Hat, I	nc. Com	mon Stoc	: k an 13-4	:00pm EST	
Open:	42.70	52wk Ra	nge:	31.77 - 52.00						43
Bid:	38.00 x 100	Volume:		1,378,416	۸.					43
Ask:	43.18 x 200	Avg Vol (3m):	2,418,720	North March	Mi		1		42
1y Target Est:	50.50	Market C	ap:	8.19B	1/w	w w	man	Num	www.	42
Beta:	0.86	P/E (ttm):		58.20	© Yahoo!					42
Next Earnings Date:	N/A	EPS (ttm):	0.73	10am	12pm		2pm Pre	4pm vious Close	
Comparison			24D+	N/A (N/A)	1d 50	d 3m	6m 1y	2у	5y ma	x
Symbol	% Cł	ng	Mkt Cap							
RHT	\$ 1.5	8%	8.19B							
MSFT	1 0.8	9%	237.64B							
ORCL	1 0.6	3%	137.41B							
SAP	† 2.4	6%	64.91B							

Business Stories

Forbes

Search news, business leaders, and stock quotes

Forbes Earnings Preview: Red Hat

By <u>Narrative Science</u>

+ Comment now

Since September 15, 2011 **Red Hat's (RHT)** stock has risen 13.9% to close at \$46.98 on December 13, 2011. Red Hat will look to keep the momentum going when it reports third quarter earnings on Monday, December 19, 2011.

What to Expect:

The average estimate is for profit of 19 cents per share, a rise of 26.7% from the company's actual earnings for the same quarter a year ago. For the fiscal year, analysts are projecting earnings of 78 cents per share.

Revenue is projected to be \$289.5 million for the quarter, 22.9% above the year-earlier total of \$235.6 million. It is expected to roll in at \$1.13 billion for the year.

Analyst Ratings:

The majority of analysts (72.7%) rate Red Hat as a buy. This compares favorably to the analyst ratings of its nearest 10 competitors, which average 63.5% buys. Analyst sentiment has been improving recently, as the number of buy ratings has risen slightly over the past three months.



Help | Login | SignUp

Trends to Watch For:

Revenue has grown by double digits year over year for the last four quarters. Over that span, the company has averaged growth of 25.2%, with the biggest boost coming in the most recent quarter when revenue rose 28% from the year earlier quarter.

The company's net income has been growing for three straight quarters. The 69% year-over-year growth in net income in the most recent quarter came after the 34.8% profit growth in the first quarter and the 43.4% rise in the fourth quarter of the last fiscal year.

Competitors:

Red Hat provides open source software solutions that customers are allowed to update, monitor and manage. One of Red Hat's main competitors in the software industry, <u>Microsoft</u> (MSFT), will report earnings on January 19, 2012. Other competitors in the information technology sector include: <u>Oracle</u> (ORCL), <u>CA</u> (CA), and <u>BMC Software</u> (BMC).

Earnings estimates provided by Zacks.

<u>Narrative Science</u>, through its proprietary artificial intelligence technology platform, turns data into stories and insights.

Another Earnings Preview

FORDES 🔻 | 'The Fault in Our Stars' Numbers Prove Young Women Are The Hottest Box Office.... Active on Facebook 🖪 💌

INVESTING 5/19/2014 @ 10:00AM 229 views

Search companies, people a



Home Depot Earnings Projected to Increase

Narrative Science

+ Comment Now + Follow Comments

By Narrative Science



full bio →

Opinions expressed by Forbes Contributors are their own.



Share

Cookies on Forbes



SageVoice: Three Questions To Ask Your Lawyer Before Starting A Small Business

CONFERENCES AND MORE

Wall Street is optimistic about **Home Depot**, which is slated to report its first quarter results on Tuesday, May 20, 2014. Analysts project a profit of 99 cents a share, a rise from 83 cents per share a year ago.

The consensus estimate has dipped over the past three months from \$1.02. For the fiscal year, analysts are projecting earnings of \$4.40 per share. Revenue is projected to be \$19.95 billion for the quarter, 4% above the year-earlier total of \$19.12 billion. For the year, revenue is projected to come in at \$82.65 billion.

A year-over-year drop in revenue in the fourth quarter ended a streak of revenue growth that had lasted three quarters.

WELCOME TO OUR WORLD



First Prototype at Northwestern

0					1		12	Stat	s Monk	ey					-				
http://jermaine.cs	.northy	veste	rn.ed	Wikim	ball/in	lex.htr	mi?path	Por	p://jer	maine.c	s.nor	hweste	ern.c	C	Q+ (loogle			
ne ranoo: Google w	laps	Tourn	ube	ткр	eura r	iews (J	130) •	Fopi	alat •										
sta		5 Pl								Selec	tag	ame d	ate:	clie	c <mark>k to s</mark>	ee cale	ndar	Þ	
			1		the D/	ATA		t	he STO	RY		the A	NALY	SIS		the	OUTL	INE	
08/12/09 White Sox <i>at</i> M	arine	ers										WRI	TE A	STO	ORY V	VITH Т	HIS D	ATA	
LINE SCORE																			
Teams		1		2	3 4	5	6	7	8	9	10	11	1	2	13	14	FIN	AL	
White Sox		0)	0 0	0 0	0	0	0	0	0	0	0		D	0	0		0	
Mariners		0)	0 (0 0	0	0	0	0	0	0	0		D	0	1		1	
BOX SCORE																			
Mariners	АВ	R	н	HR	RBI	BB	SO		Wh	te Sox		AB	R	н	HR	RBI	BB	SO	
Mike Sweeney	6	0	0	0	0	0	0		Mark	Kotsay		5	0	2	0	0	1	1	
Adrian Beltre	6	1	3	0	0	0	0		Jim	Thome		4	0	1	0	0	2	1	
Russell Branyan	5	0	1	0	0	1	2		A.J. P	ierzynsl	il	5	0	1	0	0	1	1	
Jack Wilson	1	0	0	0	0	0	0		Scott F	odsedn	ik	6	0	2	0	0	0	2	
Ichiro Suzuki	5	0	1	0	0	1	1		Ale	x Rios		6	0	1	0	0	0	2	
Ryan	4	0	0	0	0	0	1		Jay	son Nix		5	0	2	0	0	1	1	
Langernans									Carlos	Quanti		6	0	0	0	0	0	2	

First Prototype: Derived Features



First Prototype: The Outline



First Prototype: The Story





Why Stories?

12

7,622

Gross Profit Steady increase in both sales and gross profit. During this period, however, two trends have become apparent.

While sales have increased, <u>there has been a steady</u> <u>decrease in margins</u> which should be cause for concern.

In addition, while we have seen quarter-to-quarter increases in both sales and gross profit, <u>each of these</u> <u>metrics is decelerating</u>, indicating that the business is experience as a structure of the set of t

Why Stories?

- Stories *highlight*, *connect*, and *summarize* the critical aspects of a situation.
- Stories *explain*: They convey trends, causes, and even recommendations.
- Stories provide **context**.
- Stories make data **meaningful**.

Even Little Data is Big Data

C	Q1 - 2010	Q2 - 2010	Q3 - 2010	Q4 - 2010	Q1 - 2011	Q2 - 2011	Q3 - 2011	Q4 - 2011	Q1 - 2012	Q2 - 2012	Q3 - 2012
Gross Profit	\$9,600,000	\$12,402,000	\$15,774,720	\$19,750,349	\$24,334,442	\$29,497,712	\$35,168,885	\$41,229,808	\$47,513,986	\$53,809,589	\$59,867,62
Sales	\$12,000,000	\$15,600,000	\$19,968,000	\$25,159,680	\$31,198,003	\$38,061,564	\$45,673,877	\$53,895,174	\$62,518,402	\$71,270,979	\$79,823,49

This is Not a Story

In Q1, sales were \$12.6M; gross profit was \$9M. In Q2, sales were \$15.6M; gross profit was \$12.4M. In Q3, sales were \$19.97M; gross profit was \$15.77M.

Storytelling Requires Selection

- The core of storytelling is *not* NLG in a narrow sense i.e., *story-telling* != *text generation*.
 - The problem is much more *what* to say, rather than *how* to say it.
- The storyteller needs some kind of model of what matters more — of what the listener / reader *needs to know*.
 - I.e., a model of what *questions* the listener / reader has (or should have).

Stories Answer Questions

- Data storytelling revolves around *questions*.
 - A good story answers the questions that the listener / reader has (or should have).
 - The answers will often raise more questions which the story must then go on to answer.
- In other words, a (data-driven) story consists of a pragmatically grounded, logical, and structured series of *questions and answers*.
 This is what makes the story *coherent*.

• What is the nature of the *data* needed to answer the question?

- What is the nature of the *data* needed to answer the question?
- What sets of **analytics** are needed to answer the question?

- What is the nature of the *data* needed to answer the question?
- What sets of **analytics** are needed to answer the question?
- What kinds of *results* do these analytics return?

- What is the nature of the *data* needed to answer the question?
- What sets of **analytics** are needed to answer the question?
- What kinds of *results* do these analytics return?
- How can we form these results into answers to the question? How should we characterize the results?

- What is the nature of the *data* needed to answer the question?
- What sets of **analytics** are needed to answer the question?
- What kinds of *results* do these analytics return?
- How can we form these results into answers to the question? How should we characterize the results?
- What *other questions* might the answers raise?
• Data: Last quarter's sales by salesperson

- <u>Data</u>: Last quarter's sales by salesperson
- <u>Analytics</u>: *max*; or *sort* + *first*

- <u>Data</u>: Last quarter's sales by salesperson
- <u>Analytics</u>: *max*; or *sort* + *first*
- <u>Results</u>: A name and/or id, and a number; and/or a sorted list of names / ids and numbers

- <u>Data</u>: Last quarter's sales by salesperson
- <u>Analytics</u>: *max*; or *sort* + *first*
- <u>Results</u>: A name and/or id, and a number; and/or a sorted list of names / ids and numbers
- Forms of answers / characterizations of results:
 - X's sales were the maximum (at ...).
 - X had the highest sales.
 - X was the best salesperson.
 - Etc.

"Highest" != "Best"

- "Highest" (or "greatest", or "largest") can be used to characterize the results of these analytics applied to **any** metric.
- "Best" can only be used if the metric is also a *performance metric* i.e., a metric that is related to a *goal*.
 - Jupiter is not the best planet in the solar system simply because it has the highest mass.

"Highest" != "Best"

- In fact, this is important for something I glossed over at the beginning of this example — how do we get from "best salesperson" to "salesperson with the best sales" in the first place?
 - This doesn't necessarily matter in a storytelling system, but it's critical if we want a conversational interaction in natural language.
- "Highest" != "Best" is what underlies the pragmatic inference that "best X" means "X that is best in terms of some relevant performance / evaluation metric for entities of type X".

Grice's Maxims Matter

- The appropriate interpretation / characterization of the results of the data analytics must take Grice's maxims into account.
- If, e.g., there were two salespeople at the top whose sales were very close, then it would be *misleading* to simply name the top salesperson.
 A cooperative interlocutor will therefore note this in characterizing the results.
- Whether this should be an explicit inference, or simply built-in to a practical system, is debatable.

Narrative Analytics

- Narrative goals determine data analysis and even the data you need.
- Stories embody high-level patterns or themes

 such as comeback, fade, low-hanging fruit, etc. that:
 - Render data coherent.
 - Usefully **characterize** situations in broad terms.
 - Are efficiently **communicated** and effectively **grasped**.

The Story Generation Process

The Story

<u>Rhetorical structure</u>: How to organize the information.

<u>Pragmatics</u>: What the reader needs or wants to know. (The questions the story must answer.)

Semantics: The answers to the questions. (Propositional characterizations of the data and analytic results.)

<u>Analytics</u>

The Data

Civic Data

The ProPublica Nerd Blog

Secrets for Data Journalists and Newsroom Developers



How To Edit 52,000 Stories at Once

by Scott Klein ProPublica, Jan. 24, 2013, 3:19 p.m.

Today we're launching an update to our Opportunity Gap news application, with two new data points, better design on smartphones and integration with the Foursquare social network.

Also launching with this new release are short narrative descriptions of almost all of the more than 52,000 schools in our

2	school-report-010114000411.json	Jan 4, 2013 1:35 PM	2 88	JSON.	
	school-report-010114001860.json	Jan 4, 2013 1:33 PM	2 KB	JSON .	
Ĉ	school-report-010207000790.json	Jan 4, 2013 1:34 PM	2 88	JSON	
	school-report-010222000846.json	Jan 4, 2013 1:34 PM	2 88	J50N	
ľ	school-report-010265001573.json	jan 4, 2013 1 35 PM	2 8/8	J\$ON	
1	school-report-010297000152.json	jan 4, 2013 1-31 PM	2 88	JSON	
ľ	school-report-040724002267.json	Jun 4, 2013 1:32 PM	2 KB	J50N	
ē	school-report-040906000939.json	Jan 4, 2013 1:35 PM	2 8/8	J50N	
2	school-report-060001409082.json	jun 4, 2013 1-33 PM	2 88	J50N	
ē	school-report-060681000608.json	Jan 4, 2013 1:32 PM	2 KB	JSON .	
ē	school-report-060744002334.jton	jan 4, 2013 1:32 PM	2 88	JSON.	
ľ	school-report-060846011326.jton	Jan 4, 2013 1:33 PM	2 KB	J5ON	
ē	school-report-061350001542.jton	Jan 4, 2013 1:33 PM	2 8/8	JSON	
2	school-report-061488001858.json	Jan 4, 2013 1:34 PM	2 KB	JSON.	
1	school-report-061518001908.json	Jun 4, 2013 1:33 PM	2 88	JSON	
2	school-report-062019002414.json	Jan 4, 2013 1:32 PM	2 88	J50N	
2	school-report-062271002889.json	Jun 4, 2013 1-32 PM	2 88	JSON	
ē	school-report-062271002954 Json	Jan 4, 2013 1 33 PM	2 KB	JSON .	
ē	school-report-062361003588.json	Jan 4, 2013 1 34 PM	1 88	JSON.	
	school-report-062547005903.json	Jan 4, 2013 1:35 PM	2 KB	J5ON	
	school-report-062628003923.json	jan 4, 2013 1 34 PM	2 8/8	JSON.	
	actived server of STEPARATER land	114 A 2013 L 23 Dec	2.40	(CON)	

Civic Data: Education Story

Evanston Township High School

Take Chemistry

28%

1600 DODGE AVE, EVANSTON, ILL., 60204 Grades 9-12

What's the story behind this

Please share any tins or local

data?

Districts with 3,00	00 or more s	tudents
---------------------	--------------	---------

Compare Hint: Click one of t	the 🔀	Students	Total Teachers	Inexp. Teachers	AP Courses
You can compare this to compare this so	search box shool with	2,875	225	15%	23
schools in III, that are in students or more. Sele comparisons below, o		2,875	225	15%	23
		1.36M	84,195	14%	11
to nearby schools	Percentage of re	elevant students	who		State Average District Average
to low and high poverty schools or to another Illinois school	Get Free/Reduced Price Lunch	39%		Evanston Township High Illinois, is part of the Eva	n School, in Evanston, Inston Township High
Share	Take at Least One AP Course	22%		School District 202. The 2,875 students in grades has 225 teachers on staff	school reports enrolling nine through 12, and it
comparisons and comments with us and your friends.	AP Pass Rate	76%		Evanston Township High School i line with the district average in ter who qualify for free or reduced-prio	is below the state average and in ms of the percentage of its students ce lunches. On average, 44 percent
Use this hashtag to share your insights on Twitter: #myschoolyourschool	Take Advanced Math	21%		of students in Illinois qualify for free or reduced-price lunch programs, whereas 39 percent of Evanston Township High School students do. At the district level, 39 percent qualify.	

ProPublica's analysis found that all too often, states and schools provide poor students fewer educational programs like Advanced Placement, gifted and talented programs, and advanced math and science classes. Studies have linked participation in these programs

Narratives for Tableau

Tableau® has released an innovative suite of business analytics

Narratives for Qlik

Companies of all sizes use Qlik solutions to visualize and explore

Narratives for Power Bl

Microsoft Power BI is a collection of software services, apps, and connectors that work together to turn your unrelated sources of data into coherent, visually immersive, and interactive insights.

Narratives for Power BI enables Power BI users to immediately gain insight from all of their data by transforming it into intelligent narratives. These dynamic narratives automatically identify relationships in data, uncover insights in visualizations, and communicate results in natural language.

TRY IT NOW





Sales \$ by District

This analysis measures Sales \$ by District.

- Total Sales \$ is \$983.2 million across all six districts.
- The distribution ranges from \$89.3 million (District #06) to \$210.7 million (District #03), a difference of \$121.3 million.
- The average Sales \$ per district is \$163.9 million and the median is \$173.2 million.
- Sales \$ is relatively evenly distributed across all the districts.
- The top two districts represent over a quarter (41%) of overall Sales \$, and the top three districts account for over a half (60%).

SAME CH	Blue Ridge Su	dden Impact 0	vs. <u>Richmo</u>	nd Rampage 6	Scored by Cathy Guertler	Search Teams 9
Richmond	Team	1 2 3 4 5 6 7	г <mark>в н е</mark> Fina	al - League Game		Tweet Like
OFFICIAL ACTIVE	Blue Ridge Sudden I	0 0 0 0 0 0	0 1 2			
HOME SCHEDULE & RE	Richmond Rampage	1 2 0 0 3 0	6 7 1			
« All Games Recap Story						
Blue Ridge Sudd	Richmond	Rampage	escapes	Blue Ridge	Sudden	
Team 1 Blue Ridge Sudden I 0 Richmond Rampage 1	Impact bel	hind Morri	s' five str	rikeouts, 6-0		
The more than page 1	A superb effort by <u>Kat</u> Impact on Sunday at D	Morris lifted <u>Richmon</u> Darden Towe Charlotte	d Rampage to a 6-0 sville.) win in seven innings over I	Blue Ridge Sudden	
Richmond R Impact behi	Morris kept Blue Ridge Sudden Impact off balance while she was in the circle. Blue Ridge Sudden Impact managed just one hit off of <u>Morris</u> , who allowed no earned runs, walked one and struck out five during her seven innings of work.					generated from play- are currently
A superb effort by <u>Kat Mo</u> Impact on Sunday at Dard	Richmond Rampage blew out Blue Ridge Sudden Impact thanks to seven hits, including three extra base hits. Richmond Rampage scored three runs in the fifth on an RBI double by Madison Foster and a two-run triple by Haley Marshall.				ent about the accuracy echnology,	
Morris kept Blue Ridge Su managed just one hit off of	Marshall racked up thr	ree RBIs on one hit for	Richmond Rampa	ge.		
innings of work. <u>Richmond Rampage</u> blew <u>Bichmond Bampage score</u>	Richmond Rampage d choice.	idn't re <mark>linquish the l</mark> ea	d after scoring one	run in the first inning when i	it scored on a fielder's	
Haley Marshall.	Richmond Rampage added two more runs in the bottom of the second. A clutch ground out scored MoMo Bush to set the tone for the inning. That was followed up by a groundout by Erica Harris that scored Stephanie					
Richmond Rampage didn' choice.	"Powered by <u>Narrative</u> republication of this sto	Science and GameC ry must include the pr	hanger. Copyright 2 eceding attribution.	011. All rights reserved." An	ny reuse or	
Richmond Rampage adde to set the tone for the inning	. That was followed up by a	groundout by Erica Har	ris that scored <u>Stepha</u>	nie		1

(

The New Yorker



The Punch Line

6. "Benner had a good game at the plate for the Hamilton A's-Forcini. Benner went 2 for 3, drove in one, and scored one run. He singled in the third inning and doubled in the fifth inning."

This was written by a computer algorithm from Narrative Science, a company your dad thinks you might want to work for, which is why he sends you every article about how innovative the A.I.-powered language-generation platform is, even though it is literally the opposite of what you do and against everything you believe.

Conversational Interaction

- Conversational interaction for real tasks e.g., decision-making tasks.
- Pragmatically motivated common sense inference.
- Building on top of neural language models as a substrate — and understanding (or trying to) and improving their semantic competence.

Compositionality

- How "compositional" are statistical representations of language such as word embeddings?
- Definition modelling: Can we "read out" the definition of a word from its embedding?

http://thor.cs.northwestern.edu:24603

Conversational Recommendation

- User provides initial constraints.
- Concierge makes a recommendation.
- User critiques recommendation.
- Concierge understands user's needs better, makes a new recommendation.
- Loop until good enough.

Conversational Recommender

The Capital Grille



What's it about the food that you didn't like?

It seems a little tacky for the price tag

I hear you. You prefer a more stylish place.

You might want to check out Shanghai Terrace for that. Here's what a previous customer wrote: "Really amazing rooftop dining area, one of the best"

Tell me about it then Back to the options Show a new list of options

Architecture



Pragmatic Common Sense

- "I don't see any vegetarian options."
- Requires *critique* □ *preference* inference.

Preference Understanding

- Need to infer (positive) user preferences from (negative) user critiques of current offering.
 - Why? To find restaurants with attributes that match those preferences.
- Restaurant attributes are embodied in online customer reviews — which are *open-ended*.
- So we'd like to perform this inference in a similarly open-ended way.
- => Use statistical language models to perform this inference.

Preference Understanding

XLNet ("It looks too spicy for me => I prefer milder food. (...) It looks too greasy for me => I prefer lighter food. I don't eat meat => I prefer "

→ "salad."

XLNet ("(...) I don't like fast food => I prefer ")

→ "healthy cuisine."

XLNet ("(...) I don't eat fish => I prefer ")

→ "chicken, pork, lamb, duck or beef."

Matching Preferences to Reviews

preferences = ["I prefer to socialize over a meal.", "I prefer light fish, like salmon."] reviews = $[review_1, review_2, review_3, review_4, review_5]$ sentences = parse sentences(reviews) positive sents = filter positive(sentences, threshold=0.80) ["Been here half dozen times over the last several years.", "Always good service, always good food and always delicious.", When you walk in, there's a half dozen or so little private booths behind curtains to the far left - ask for one of those." "It's kind of cozy.", "My friend always gets the lobster bisque (it's pretty and they pour it over while you watch).", "I always get something seafoody and it's always delicious.", "Will be back."(...)] for each pref in preferences: similar sents = rank semantic similarity(pref, positive sents) arguments.add(max(similar sents)) return max(arguments) arguments = [(0.59, "Anyway, this fancy restaurant is the perfect place for people and families who want to spend good time with each other and quality foods."), (0.61, "Ceder plank salmon was good.")]

Language Models for Common Sense Inference

- We're using language models directly to go from user critiques to user preferences.
- We're using something else to go from user preferences to restaurant recommendations.
- Why? Crossing the *pragmatic gap*.
 - The language model doesn't "know" that the goal / task is to choose a restaurant.
 - I.e., how do we go from "M & S has great seafood" to "If you like seafood, pick M & S"?

Two Possible Approaches

- Use language models (more generally, statistical machine learning) configured to carry out specific sub-tasks as *components* of larger (heterogeneous) systems (as we did above).
 This is how self-driving cars work.
- Figure out ways to get such models to *internalize* these sub-tasks — i.e., scaffold them through training or architecture to directly carry out these sub-tasks, and then to *compose these properly to carry out the task as whole*.

Decision-Making as a Task

- System needs to know the *facts* (Miami has an average temperature of 75 F, Minneapolis of 50 F).
- It needs to *compare* (Miami is generally warmer than Minneapolis).
- It needs to know or learn *user preference* (I prefer warmer weather).
- It needs to combine these to make a decision / recommendation (I should go to Miami).

Compositional Fine-Tuning

- Decompose the task into component inferences (as illustrated above).
- Explicitly fine-tune a language model on these inferences.
- It works better than a recent similar approach ("chain of thought" inference) — and it works *much* better when using smaller (and cheaper) models.
- (When does it make sense to use this approach?)

Decision Model



Results

Base Model	Method	Score on Decision Templates
DaVinci	8-Shot Prompting	0.83 ± 0.08
DaVinci	8-Shot Chain of Thought	0.98 ± 0.02
Curie	8-Shot Chain of Thought	0.50 ± 0.12
Curie	CFT on Factual Statements, Factual Comparisons, and Decision Templates	0.95 ± 0.01

Base Model	Method	Score on Decision Templates
DaVinci	8-Shot Prompting	0.54 ± 0.09
DaVinci	8-Shot Chain of Thought	0.55 ± 0.07
Curie	8-Shot Chain of Thought	0.50 ± 0.06
Curie	CFT on Factual Statements, Factual Comparisons, and Decision Templates	0.74 ± 0.05

Table 1: Comparison to chain of thought in the world travel domain (above) and local dining (below). CFT performs as well as chain of thought prompting for world cities and 35% better for local restaurants, with an LM only 7.4% of the size (13B vs 175B).

We Don't "See" the World the Way Computers Do

"Robin"



"Speed Limit 45"



Two Big Opportunities

- How to make computers more like us.
- How to build *collaborative human-AI systems* that build on the strengths of each to produce outcomes that either alone aren't able to achieve.

PS I Am Looking For Students!

- Please send me a note if you're interested:
- <u>l-birnbaum@northwestern.edu</u>

Thanks!

- Colleagues (many), especially Doug Downey and Kris Hammond.
- Students (scores), especially Nick Allen, Jay Budzik, Victor Bursztyn, Miriam Boon, Jason Cohn, Francisco Iacobelli, Jiahui Liu, Nate Nichols, Shawn O'Banion, John Templon, and Earl Wagner.
- Funders, especially the John S. and James L. Knight Foundation, the National Science Foundation, Google, and Adobe.
- Narrative Science (now part of Salesforce and its Tableau unit), for making it real.